

COMMUTATORS OF TRACE ZERO MATRICES OVER PRINCIPAL IDEAL RINGS

ALEXANDER STASINSKI

ABSTRACT. We prove that for every trace zero square matrix A of size at least 3 over a principal ideal ring R , there exist trace zero matrices X, Y over R such that $XY - YX = A$. Moreover, we show that X can be taken to be regular mod every maximal ideal of R . This strengthens our earlier result that A is a commutator of two matrices (not necessarily of trace zero), and in addition, the present proof is simpler than the earlier one.

Shalev has conjectured an analogous statement for group commutators in SL_n over p -adic integers. We prove Shalev's conjecture for $n = 2$.

1. INTRODUCTION

Let R be a principal ideal ring, which we will always take to be commutative with identity (e.g., R could be a field). We let $\mathfrak{gl}_n(R)$ denote the Lie algebra of $n \times n$ matrices over R with Lie bracket $[X, Y] = XY - YX$, and $\mathfrak{sl}_n(R)$ the sub Lie algebra of trace zero matrices. In case $R = K$ is a field, a theorem of Albert and Muckenhoupt [1] says that every $A \in \mathfrak{sl}_n(K)$ is a commutator in $\mathfrak{gl}_n(K)$, that is, there exist $X, Y \in \mathfrak{gl}_n(K)$ such that $[X, Y] = A$. To go beyond the field case requires new ideas and the first major step was taken by Laffey and Reams [4] who proved the analogous result for $R = \mathbb{Z}$, solving a problem posed by Vaserstein [11, Section 5]. Whether every element in $\mathfrak{sl}_n(R)$ is a commutator in $\mathfrak{gl}_n(R)$ for a PIR R , was an open problem going back implicitly at least to Lissner [5], and was settled in the affirmative in [9].

In light of the above results, a natural question is whether X and Y can be taken in $\mathfrak{sl}_n(R)$, rather than just $\mathfrak{gl}_n(R)$. When $R = K$ is a field, it is known by work of Thompson [10, Theorems 1-4] that any $A \in \mathfrak{sl}_n(K)$ can be written as $A = [X, Y]$ for some $X, Y \in \mathfrak{sl}_n(K)$, except when $\text{char } K = 2$ and $n = 2$. A generalisation of Thompson's result, allowing X and Y to lie in an arbitrary hyperplane in $\mathfrak{gl}_n(K)$ (but assuming $n > 2$ and $|K| > 3$), was recently obtained by de Seguins Pazzis [6]. On the other hand, it does not seem possible to modify our proof in [9] to yield the stronger assertion that every $A \in \mathfrak{sl}_n(R)$, with $n \geq 3$, is a commutator of matrices in $\mathfrak{sl}_n(R)$, even in the case where R is a field.

The main result of the present paper is that for any principal ideal domain (henceforth PID) R and $A \in \mathfrak{sl}_n(R)$, with $n \geq 3$, there exist $X, Y \in \mathfrak{sl}_n(R)$ such that $A = [X, Y]$. It is also easy to see that when 2 is invertible in R , the same conclusion holds for $A \in \mathfrak{sl}_2(R)$. Moreover, it follows from our proof that X can be chosen to be regular mod every maximal ideal of R (this was stated as an open problem in [9]). Our proof is significantly simpler than the proof of the main result

in [9], and the new idea is to consider the matrices

$$X(\mathbf{x}, a) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ x_1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ \vdots & 0 & \vdots & \ddots & 1 \\ x_{n-1} & a & 0 & \cdots & 0 \end{pmatrix} \in \mathfrak{sl}_n(R),$$

where $\mathbf{x} = (x_1, \dots, x_{n-1})^\top \in R^{n-1}$ and $a \in R$; see Section 3. These matrices have some remarkable properties which let us carry through the proof. More precisely, we show that for a given non-scalar $A \in \mathfrak{sl}_n(R)$ in Laffey–Reams form (see [9, Theorem 5.6]), we can find \mathbf{x} and a such that

$$\mathrm{tr}(X(\mathbf{x}, a)^r A) = 0, \quad \text{for } r = 1, \dots, n-1,$$

and at the same time ensure that $X(\mathbf{x}, a) \bmod \mathfrak{p}$ is regular in $\mathfrak{gl}_n(R/\mathfrak{p})$, for every maximal ideal \mathfrak{p} of R , as well as regular in $\mathfrak{sl}_n(R/\mathfrak{p})$, for any \mathfrak{p} for which A is non-scalar mod \mathfrak{p} . We note that the condition on the vanishing of traces above is rather delicate, given that we also want $X(\mathbf{x}, a)$ to have the above regularity property and trace zero, and depends on the existence of a solution of a system of polynomial equations over R , which in most cases is hopelessly complicated. Nevertheless, for the matrices $X(\mathbf{x}, a)$ the system of equations becomes atypically simple, and we are able to show that a solution exists. We then use the well known local-global principle for systems of linear equations over rings, applied to the system defined by $[X(\mathbf{x}, a), Y] = A$, $Y \in \mathfrak{sl}_n(R)$. Working over the localisation $R_{\mathfrak{p}}$ at a maximal ideal \mathfrak{p} of R , we use a variant of the criterion of Laffey and Reams (see Section 2, Proposition 2.4) to show that the system has a solution if A is non-scalar mod \mathfrak{p} . Here we use that $A \bmod \mathfrak{p}$ is not merely regular in $\mathfrak{gl}_n(R/\mathfrak{p})$ but also regular in $\mathfrak{sl}_n(R/\mathfrak{p})$. The existence of a solution over $R_{\mathfrak{p}}$ when \mathfrak{p} is such that $A \bmod \mathfrak{p}$ is scalar is more subtle and requires a separate argument. The existence of a local solution for every maximal ideal \mathfrak{p} then implies the existence of global solution, and since any non-scalar matrix is $\mathrm{GL}_n(R)$ -conjugate to one in Laffey–Reams form, our main result follows (the case when A is scalar requires a separate discussion, but is easy).

Once the main result has been established for a PID, it is easy to deduce it for an arbitrary principal ideal ring (not necessary and integral domain).

In [7, Conjecture 1.3] and [8, Conjecture 8.10] Shalev conjectured that if either $n \geq 3$ or $n = 2$ and $p > 3$, then any element in $\mathrm{SL}_n(\mathbb{Z}_p)$ is a commutator of elements in $\mathrm{SL}_n(\mathbb{Z}_p)$ (here \mathbb{Z}_p denotes the p -adic integers). The main result of the present paper, in the case where $R = \mathbb{Z}_p$, is therefore a Lie algebra analogue of this conjecture. In [2] some results towards this conjecture were obtained. In Section 6 we give a proof of Shalev’s conjecture for $n = 2$.

We end this introduction with a word on notation. A ring (without further specification) will mean a commutative ring with identity. Throughout, we will use 1_n to denote the identity matrix in $\mathfrak{gl}_n(S)$, where S is a ring with identity. If $X \in \mathfrak{gl}_n(S)$, $S[X]$ will denote the unital S -algebra generated by X .

2. THE CRITERION OF LAFFEY AND REAMS

In this section, K denotes an arbitrary field. We will prove an analogue of the Laffey–Reams criterion (see [4, Section 3] and [9, Proposition 3.3]) for a matrix in

$\mathfrak{sl}_n(R)$, R a local PID, to be a commutator of matrices in $\mathfrak{sl}_n(R)$. This criterion plays a key role in our proof of the main theorem.

We need a couple of remarks about regular elements in $\mathfrak{sl}_n(K)$. It is well known that an element $X \in \mathfrak{gl}_n(K)$ is regular if and only if

$$C_{\mathfrak{gl}_n(K)}(X) = K[X],$$

that is, if and only if the centraliser of X in $\mathfrak{gl}_n(K)$ has dimension n . In this situation, we will say that X is $\mathfrak{gl}_n(K)$ -regular. Similarly, if $X \in \mathfrak{sl}_n(K)$ we define X to be $\mathfrak{sl}_n(K)$ -regular if

$$\dim C_{\mathfrak{sl}_n(K)}(X) = n - 1.$$

For $X \in \mathfrak{sl}_n(K)$ it may happen that X is $\mathfrak{gl}_n(K)$ -regular but not $\mathfrak{sl}_n(K)$ -regular: take for example $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in \mathfrak{sl}_n(\mathbb{F}_2)$.

The following result describes the precise relationship between the properties \mathfrak{sl}_n -regular and \mathfrak{gl}_n -regular over a field.

Lemma 2.1. *Let $X \in \mathfrak{sl}_n(K)$. Then the following holds:*

- (i) *If X is $\mathfrak{sl}_n(K)$ -regular, then X is $\mathfrak{gl}_n(K)$ -regular.*
- (ii) *X is $\mathfrak{sl}_n(K)$ -regular if and only if it is $\mathfrak{gl}_n(K)$ -regular and $\text{tr}(K[X]) \neq 0$.*
- (iii) *If $\text{char } K$ does not divide n , then an element X is $\mathfrak{sl}_n(K)$ -regular if and only if it is $\mathfrak{gl}_n(K)$ -regular.*

Proof. For the first part, note that $C_{\mathfrak{sl}_n(K)}(X)$ is either equal to $C_{\mathfrak{gl}_n(K)}(X)$ or is a hypersurface in $C_{\mathfrak{gl}_n(K)}(X)$, so $C_{\mathfrak{sl}_n(K)}(X)$ has codimension at most one in $C_{\mathfrak{gl}_n(K)}(X)$. Thus X being $\mathfrak{sl}_n(K)$ -regular implies that $\dim C_{\mathfrak{gl}_n(K)}(X) \leq n$. But it is well-known that the dimension of a centraliser in $\mathfrak{gl}_n(K)$ is always at least n , so X is $\mathfrak{gl}_n(K)$ -regular.

For the second part, first note that $C_{\mathfrak{sl}_n(K)}(X)$ is the kernel of the trace map $\text{tr} : C_{\mathfrak{gl}_n(K)}(X) \rightarrow K$. Now, if X is $\mathfrak{sl}_n(K)$ -regular, then by the previous part, X is $\mathfrak{gl}_n(K)$ -regular, so $C_{\mathfrak{gl}_n(K)}(X) = K[X]$. Thus $\dim C_{\mathfrak{sl}_n(K)}(X) = n - 1$ implies that this trace map is surjective, that is, that $\text{tr}(K[X]) \neq 0$. Conversely, if X is $\mathfrak{gl}_n(K)$ -regular and $\text{tr}(K[X]) \neq 0$, then $\dim C_{\mathfrak{gl}_n(K)}(X) = n$ and $\text{tr} : C_{\mathfrak{gl}_n(K)}(X) \rightarrow K$ is surjective, so the kernel has dimension $n - 1$.

Finally, when $\text{char } K$ does not divide n and X is $\mathfrak{gl}_n(K)$ -regular, then $\text{tr}(1_n) = n \neq 0$, so the previous part implies that X is $\mathfrak{sl}_n(K)$ -regular. \square

Proposition 2.2. *Let $X \in \mathfrak{sl}_n(K)$ be $\mathfrak{sl}_n(K)$ -regular and let $A \in \mathfrak{sl}_n(K)$. Then $A = [X, Y]$ for some $Y \in \mathfrak{sl}_n(K)$ if and only if $\text{tr}(X^r A) = 0$ for all $r = 1, \dots, n - 1$.*

Proof. Since X is $\mathfrak{gl}_n(K)$ -regular by Lemma 2.1, the set $\{1_n, X, \dots, X^{n-1}\}$ is linearly independent over K , so the subspace

$$V = \{B \in \mathfrak{sl}_n(K) \mid \text{tr}(X^r B) = 0 \text{ for } r = 1, \dots, n - 1\}$$

has dimension $n^2 - n$. The kernel of the linear map $\mathfrak{sl}_n(K) \rightarrow \mathfrak{sl}_n(K)$, $Y \mapsto [X, Y]$ is equal to the centraliser $C_{\mathfrak{sl}_n(K)}(X)$, which has dimension $n - 1$ since X is $\mathfrak{sl}_n(K)$ -regular. Thus the image $[X, \mathfrak{sl}_n(K)]$ of the map $Y \mapsto [X, Y]$ has dimension $n^2 - n$. But if $A \in [X, \mathfrak{sl}_n(K)]$, there exists a $Y \in \mathfrak{sl}_n(K)$ such that for every $r = 1, \dots, n - 1$ we have

$$\text{tr}(X^r A) = \text{tr}(X^r (XY - YX)) = \text{tr}(X^{r+1} Y) - \text{tr}(X^r Y X) = 0.$$

Thus $[X, \mathfrak{sl}_n(K)] \subseteq V$. Since $\dim V = \dim[X, \mathfrak{sl}_n(K)]$ we conclude that $V = [X, \mathfrak{sl}_n(K)]$. \square

If S is a ring, $I \subseteq S$ an ideal and $X \in \mathfrak{gl}_n(S)$, we denote by X_I the image of X under the canonical map $\mathfrak{gl}_n(S) \rightarrow \mathfrak{gl}_n(S/I)$.

Lemma 2.3. *Let S be a local commutative ring (with identity) with maximal ideal \mathfrak{m} . Let $X \in \mathfrak{sl}_n(S)$ be such that $X_{\mathfrak{m}}$ is $\mathfrak{sl}_n(S/\mathfrak{m})$ -regular. Then the canonical map*

$$C_{\mathfrak{sl}_n(S)}(X) \longrightarrow C_{\mathfrak{sl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}})$$

is surjective.

Proof. As $C_{\mathfrak{sl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}})$ has dimension $n - 1$ and is the kernel of the trace map $\text{tr} : C_{\mathfrak{gl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}}) \rightarrow S/\mathfrak{m}$, this map must be surjective. Thus, there exists an $a \in C_{\mathfrak{gl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}})$ such that $\text{tr}(a) = 1$. Since $X_{\mathfrak{m}}$ is $\mathfrak{sl}_n(S/\mathfrak{m})$ -regular, it is also $\mathfrak{gl}_n(S/\mathfrak{m})$ -regular, so

$$C_{\mathfrak{gl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}}) = (S/\mathfrak{m})[X_{\mathfrak{p}}].$$

Let $\hat{a} \in S[X]$ be any lift of a . Then $\text{tr}(\hat{a}) \in 1 + \mathfrak{m}$, so $\text{tr}(\hat{a})$ is a unit since S is a local ring. Since $S[X] \subseteq C_{\mathfrak{gl}_n(S)}(X)$, we conclude that the trace map $\text{tr} : C_{\mathfrak{gl}_n(S)}(X) \rightarrow S$ is surjective. The surjectivity of the map in the lemma is now a formal consequence. Indeed, let $b \in C_{\mathfrak{sl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}}) \subseteq (S/\mathfrak{m})[X]$, and take a lift $\hat{b} \in S[X]$ of b . Then $\text{tr}(\hat{b}) \in \mathfrak{m}$, so there exists a $c \in \mathfrak{m}C_{\mathfrak{gl}_n(S)}(X)$ such that $\text{tr}(c) = \text{tr}(\hat{b})$ (namely let $c = \text{tr}(\hat{b})c'$, where $\text{tr}(c') = 1$). Then $\hat{b} - c \in C_{\mathfrak{sl}_n(S)}(X)$, and the image of $\hat{b} - c$ in $C_{\mathfrak{sl}_n(S/\mathfrak{m})}(X_{\mathfrak{m}})$ is b . \square

The following result is a local version of the criterion of Laffey and Reams ([9, Proposition 3.3]), with the difference that we need $X_{\mathfrak{p}}$ to be $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular to ensure that $Y \in \mathfrak{sl}_n(R)$ rather than just in $\mathfrak{gl}_n(R)$.

Proposition 2.4. *Assume that R is a local PID with maximal ideal \mathfrak{p} , let $A \in \mathfrak{sl}_n(R)$ and let $X \in \mathfrak{sl}_n(R)$ be such that $X_{\mathfrak{p}}$ is $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular. Then $A = [X, Y]$ for some $Y \in \mathfrak{sl}_n(R)$ if and only if $\text{tr}(X^r A) = 0$ for $r = 1, \dots, n - 1$.*

Proof. Clearly the condition $\text{tr}(X^r A) = 0$ for all $r \geq 1$ is necessary for A to be of the form $[X, Y]$ with $Y \in \mathfrak{sl}_n(R)$. Conversely, suppose that $\text{tr}(X^r A) = 0$ for $r = 1, \dots, n - 1$. We claim that X is $\mathfrak{sl}_n(F)$ -regular, considered as an element of $\mathfrak{sl}_n(F)$. Indeed, by [9, Proposition 2.6] X is $\mathfrak{gl}_n(F)$ -regular, and since $X_{\mathfrak{p}}$ is $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular, there exists an element $a \in R[X]$ such that $\text{tr}(a) \neq 0$. Thus $\text{tr}(F[X]) \neq 0$, and so X is $\mathfrak{sl}_n(F)$ -regular by Lemma 2.1.

Now, by Proposition 2.2 we have $A = [X, M]$ for some $M \in \mathfrak{sl}_n(F)$. Let p be a generator of \mathfrak{p} . Then there exists a non-negative integer m such that $p^m M \in \mathfrak{sl}_n(R)$, and we have $[X, p^m M] = p^m [X, M] = p^m A$. Choose m to be minimal with respect to the property that $[X, C] = p^m A$ for some $C \in \mathfrak{sl}_n(R)$. Assume that $m > 0$. Then $[X_{\mathfrak{p}}, C_{\mathfrak{p}}] = 0$, so $X_{\mathfrak{p}}$ commutes with $C_{\mathfrak{p}}$. Since $X_{\mathfrak{p}}$ is $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular, there exists a $\hat{C} \in C_{\mathfrak{sl}_n(R)}(X)$ such that $\hat{C}_{\mathfrak{p}} = C_{\mathfrak{p}}$, by Lemma 2.3. Thus $C = \hat{C} + pD$, for some $D \in \mathfrak{sl}_n(R)$, so

$$[X, C] = [X, pD] = p[X, D] = p^m A.$$

Cancelling a factor of p , we obtain a contradiction to the minimality of m . Thus $m = 0$, and the result is proved. \square

3. THE MATRICES $X(\mathbf{x}, a)$

Let S be a ring (commutative with identity), $n \geq 3$, $\mathbf{x} = (x_1, \dots, x_{n-1})^\top \in S^{n-1}$ and $a \in S$. The key to our main result is to consider the following matrices:

$$X(\mathbf{x}, a) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ x_1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ \vdots & 0 & \vdots & \ddots & 1 \\ x_{n-1} & a & 0 & \cdots & 0 \end{pmatrix} \in \mathfrak{sl}_n(S),$$

that is, $X(\mathbf{x}, a) = (m_{ij})$, where

$$\begin{cases} m_{i,i+1} = 1 & \text{for } i = 2, \dots, n-1, \\ m_{i1} = x_{i-1} & \text{for } i = 2, \dots, n-2, \\ m_{n,2} = a \\ m_{ij} = 0 & \text{otherwise.} \end{cases}$$

We can write $X(\mathbf{x}, a)$ in block form as

$$X(\mathbf{x}, a) = \begin{pmatrix} 0 & \bar{0} \\ \mathbf{x} & P \end{pmatrix},$$

where $\bar{0} = (0, \dots, 0)$ is a $1 \times n$ matrix and $P = (p_{ij})$, $1 \leq i, j \leq n-1$, where $p_{i,i+1} = 1$ for $i = 1, \dots, n-2$, $p_{n-1,1} = a$ and $p_{ij} = 0$ otherwise. Thus, P is the (row-wise) companion matrix of the polynomial $x^{n-1} - a$.

Lemma 3.1. *Let $P \in \mathfrak{sl}_{n-1}(S)$ be as above, and let $\mathbf{y} = (y_1, \dots, y_{n-1})^\top \in S^{n-1}$. Then, for any $z \in S$, and $r = 1, \dots, n-1$, we have*

$$\text{tr}(P^{r-1} \mathbf{y}(z, 0, \dots, 0)) = zy_r.$$

Proof. Write $P^{r-1} = (p_{ij}^{(r-1)})$, for $1 \leq i, j \leq n-1$. Since each column in $\mathbf{y}(z, 0, \dots, 0)$, except for the first one, is zero, we have

$$\text{tr}(P^{r-1} \mathbf{y}(z, 0, \dots, 0)) = (p_{11}^{(r-1)}, p_{12}^{(r-1)}, \dots, p_{1,n-1}^{(r-1)}) z \mathbf{y}.$$

Since P is a companion matrix, there exists a $v \in S^{n-1}$ such that $\{v, Pv, \dots, P^{n-2}v\}$ is an S -basis for S^{n-1} and P is the matrix of the linear map defined by P with respect to this basis. Thus, for each $r = 1, \dots, n-1$, the first row of P^{r-1} is $(p_{11}^{(r-1)}, p_{12}^{(r-1)}, \dots, p_{1,n-1}^{(r-1)})$, where $p_{1r}^{(r-1)} = 1$ and all other $p_{1j} = 0$. Hence

$$(p_{11}^{(r-1)}, p_{12}^{(r-1)}, \dots, p_{1,n-1}^{(r-1)}) z \mathbf{y} = zy_r,$$

and the lemma follows. \square

Lemma 3.2. *For $r = 1, \dots, n-1$ we have*

$$X(\mathbf{x}, a)^r = \begin{pmatrix} 0 & \bar{0} \\ P^{r-1} \mathbf{x} & P^r \end{pmatrix},$$

In particular, $\text{tr}(X(\mathbf{x}, a)^r) = 0$ for $r = 1, \dots, n-2$, and $\text{tr}(X(\mathbf{x}, a)^{n-1}) = (n-1)a$.

Proof. The expression for $X(\mathbf{x}, a)^r$ follows easily, using block-multiplication of matrices. The assertion about the trace of $X(\mathbf{x}, a)^r$ for $r = 1, \dots, n-2$ follows from a simple induction argument, proving that for each $r = 1, \dots, n-2$, we have $P^r = (p_{ij}^{(r)})$, where $p_{i,i+r}^{(r)} = 1$ for $i = 1, \dots, n-1-r$ and $p_{n-1-r+j,j}^{(r)} = a$ for $j = 1, \dots, r$, and $p_{ij}^{(r)} = 0$ otherwise. Finally, the relation $\text{tr}(X(\mathbf{x}, a)^{n-1}) = (n-1)a$ follows from the fact that the characteristic polynomial of P is $x^{n-1} - a$. \square

Lemma 3.3. *Let K be a field, $x_1, \dots, x_{n-1} \in K^{n-1}$ and $a \in K$. If either $x_{n-1} \neq 0$ or $a \neq 0$, then $X(\mathbf{x}, a)$ is $\mathfrak{gl}_n(K)$ -regular. If $a \neq 0$, then $X(\mathbf{x}, a)$ is $\mathfrak{sl}_n(K)$ -regular.*

Proof. For simplicity, write $X = X(\mathbf{x}, a)$. We will show that if $x_{n-1} \neq 0$ or $a \neq 0$, then X is $\mathfrak{gl}_n(K)$ -regular, by showing that $\{1_n, X, \dots, X^{n-1}\}$ is linearly independent. Lemma 3.2 implies that $\{1_n, X, \dots, X^{n-2}\}$ is linearly independent because P is regular, so $\{1_{n-1}, P, \dots, P^{n-2}\}$ is linearly independent. Moreover, by Lemma 3.2 and its proof, we have

$$X^{n-1} = \begin{pmatrix} 0 & \bar{0} \\ P^{n-2}\mathbf{x} & a1_{n-1} \end{pmatrix}, \quad \text{where} \quad P^{n-2}\mathbf{x} = \begin{pmatrix} x_{n-1} \\ ax_1 \\ \vdots \\ ax_{n-2} \end{pmatrix}.$$

Thus, since P^i has zero diagonal for all $r = 1, \dots, n-2$ (see the proof of Lemma 3.2), we conclude that X^{n-1} is not a linear combination of $1_n, X, \dots, X^{n-2}$ if $a \neq 0$. On the other hand, if $a = 0$ and $x_{n-1} \neq 0$, then X^{n-1} is the matrix whose $(2, 1)$ -entry is x_{n-1} and all other entries are zero. Since each matrix in $\{1_n, X, \dots, X^{n-2}\}$ has a non-zero (i, j) -entry for some $(i, j) \neq (2, 1)$, we conclude that X^{n-1} is not a linear combination of $1_n, X, \dots, X^{n-2}$ if $a = 0$ and $x_{n-1} \neq 0$.

Suppose now that $a \neq 0$; then X is $\mathfrak{gl}_n(K)$ -regular. If $\text{char } K \nmid n$, Lemma 2.1 implies that X is $\mathfrak{sl}_n(K)$ -regular. On the other hand, if $\text{char } K \mid n$, then

$$\text{tr}(X^{n-1}) = (n-1)a = -a,$$

by Lemma 3.2, so $\text{tr}(K[X]) \neq 0$ and Lemma 2.1 implies that X is $\mathfrak{sl}_n(K)$ -regular. \square

4. THE FIELD CASE

In this section we give a proof of our main result in the case where $R = K$ is a field. We give a separate proof in this case, as it is simpler than for a general PID. The result over a field was first proved by Thompson [10], who also showed that, apart for some small exceptions, one of the matrices X can in fact be taken to be nilpotent. We give a new proof of Thompson's result, but instead of showing that X can be chosen to be nilpotent, we show that it can be taken to be $\mathfrak{gl}_n(K)$ -regular (and often $\mathfrak{sl}_n(K)$ -regular).

First let $n = 2$. For $x, y, z, s, t, u \in K$ we have

$$\left[\begin{pmatrix} x & y \\ z & -x \end{pmatrix}, \begin{pmatrix} s & t \\ u & -s \end{pmatrix} \right] = \begin{pmatrix} uy - tz & 2(tx - sy) \\ 2(sz - ux) & tz - uy \end{pmatrix}.$$

Thus, if $\text{char } K = 2$, a matrix in $\mathfrak{sl}_2(K)$ is of the form $[X, Y]$ for $X, Y \in \mathfrak{sl}_2(K)$ if and only if it is scalar. On the other hand, if $\text{char } K \geq 3$ and $a, b, c \in K$, then

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{cases} \left[\begin{pmatrix} 0 & 1 \\ -\frac{c}{b} & 0 \end{pmatrix}, \begin{pmatrix} -\frac{b}{2} & 0 \\ a & \frac{b}{2} \end{pmatrix} \right] & \text{if } b \neq 0, \\ \left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} \frac{c}{2} & -a \\ 0 & -\frac{c}{2} \end{pmatrix} \right] & \text{if } b = 0. \end{cases}$$

Note that all of the matrices involved in the above commutators are $\mathfrak{gl}_n(K)$ -regular.

Lemma 4.1. *Let S be a ring (commutative with identity) such that $n = 1 + \dots + 1 = 0$ in S . Then, for every $\lambda \in S$ there exist $X, Y \in \mathfrak{sl}_n(S)$ such that X is $\mathfrak{gl}_n(S)$ -regular and $[X, Y] = \lambda 1_n$.*

Proof. Take $X = (x_{ij})$, where $x_{i,i+1} = 1$ for $i = 1, \dots, n-1$ and $x_{ij} = 0$ otherwise, and $Y = (y_{ij})$, where $y_{j+1,j} = j$, for $j = 1, \dots, n-1$ and $y_{ij} = 0$ otherwise. Then X is a companion matrix, hence regular as an element of $\mathfrak{gl}_n(S)$. A direct computation shows that $[X, Y] = 1_n$, because $-(n-1) = 1$ in S , and thus $[X, \lambda Y] = \lambda 1_n$. \square

Remark 4.2. If $S = K$ is a field, Lemma 4.1 does not hold if X is required to be $\mathfrak{sl}_n(K)$ -regular; in fact, the X in the lemma is necessarily not $\mathfrak{sl}_n(K)$ -regular, unless $\lambda = 0$. The author was alerted to the following simple argument by a referee: Suppose that $[X, Y] = \lambda 1_n$ where $\lambda \neq 0$ and X is $\mathfrak{gl}_n(K)$ -regular. Then $\text{tr}(X^i \lambda 1_n) = \lambda \text{tr}(X^i) = 0$, hence $\text{tr}(X^i) = 0$, for all $i = 0, \dots, n-1$. Thus X is not $\mathfrak{sl}_n(K)$ -regular, by Lemma 2.1.

Theorem 4.3. *Let K be a field and $A \in \mathfrak{sl}_n(K)$, with $n \geq 3$. Then there exist $X, Y \in \mathfrak{sl}_n(K)$ such that $[X, Y] = A$. Moreover, if A is scalar, X can be chosen to be $\mathfrak{gl}_n(K)$ -regular and if A is non-scalar, X can be chosen to be $\mathfrak{sl}_n(K)$ -regular.*

Proof. Assume first that A is scalar. By [9, Proposition 4.1], there exist $X, Y \in \mathfrak{gl}_n(K)$ with X $\mathfrak{gl}_n(K)$ -regular, such that $[X, Y] = A$. If $\text{char } K$ does not divide n , then $[X - \frac{\text{tr}(X)}{n} 1_n, Y - \frac{\text{tr}(Y)}{n} 1_n] = [X, Y] = A$, where $X - \frac{\text{tr}(X)}{n} 1_n, Y - \frac{\text{tr}(Y)}{n} 1_n \in \mathfrak{sl}_n(K)$ and $X - \frac{\text{tr}(X)}{n} 1_n$ is $\mathfrak{gl}_n(K)$ -regular. On the other hand, if $\text{char } K$ divides n , then the desired assertion follows from Lemma 4.1.

Assume now that A is not scalar and let $A = (a_{ij})$. Then the rational canonical form implies that after a possible $\text{GL}_n(K)$ -conjugation, we can assume that $a_{11} = 0$, $a_{12} = 1$ and $a_{ij} = 0$ whenever $j \geq i + 2$. We will show that $x_1, \dots, x_{n-1} \in K$ can be chosen such that $\text{tr}(X(\mathbf{x}, 1)^r A) = 0$ for each $r = 1, \dots, n-1$. By Lemma 3.2 we have

$$X(\mathbf{x}, 1)^r = \begin{pmatrix} 0 & \bar{0} \\ P^{r-1}\mathbf{x} & P^r \end{pmatrix},$$

where $P = (p_{ij})$, $1 \leq i, j \leq n-1$ is such that $p_{i,i+1} = 1$ for $i = 1, \dots, n-2$, $p_{n-1,1} = 1$ and $p_{ij} = 0$ otherwise. Writing A in block-form, we have

$$A = \begin{pmatrix} 0 & (1, 0, \dots, 0) \\ \mathbf{a} & Q \end{pmatrix},$$

where \mathbf{a} is an $n \times 1$ matrix and $Q \in \mathfrak{gl}_{n-1}(K)$. Thus

$$X(\mathbf{x}, 1)^r A = \begin{pmatrix} 0 & \bar{0} \\ P^r \mathbf{a} & Q' \end{pmatrix},$$

where $Q' = P^{r-1}\mathbf{x}(1, 0, \dots, 0) + P^r Q$. Thus, by Lemma 3.1,

$$\text{tr}(X(\mathbf{x}, 1)^r A) = \text{tr}(Q') = x_r + \text{tr}(P^r Q),$$

for each $r = 1, \dots, n-1$. Put $x_r = -\text{tr}(P^r Q)$, so that $\text{tr}(X(\mathbf{x}, 1)^r A) = 0$, for $r = 1, \dots, n-1$. By Lemma 3.3 $X(\mathbf{x}, 1)$ is $\mathfrak{sl}_n(K)$ -regular, so Proposition 2.2 implies that there exists a $Y \in \mathfrak{sl}_n(K)$ such that

$$[X(\mathbf{x}, 1), Y] = A.$$

□

Remark 4.4. Our approach cannot be modified to yield Thompson's result that X can be taken to be nilpotent. The reason for this is that $X(\mathbf{x}, a)$ is nilpotent if and only if P is nilpotent if and only if $a = 0$. Therefore, even if $X(\mathbf{x}, a)$ is nilpotent and $\mathfrak{gl}_n(K)$ -regular, it cannot be $\mathfrak{sl}_n(K)$ -regular, because $\text{tr}(X(\mathbf{x}, 0)^r) = 0$ for every $r = 1, \dots, n-1$.

5. PROOF OF THE MAIN THEOREM

Throughout this section, R is an arbitrary PID with fraction field F . Note that we consider fields as special types of PIDs.

Before proving our main result (Theorem 5.3 below), we give a new and simplified proof of the main result in [9] that any $A \in \mathfrak{sl}_n(R)$ is a commutator of matrices in $\mathfrak{gl}_n(R)$. The proof of our main result is a bit harder, as it involves a special analysis for certain prime ideals. Both proofs make essential use of the Laffey-Reams form and rely on the following key result:

Lemma 5.1. *Suppose that $A = (a_{ij}) \in \mathfrak{sl}_n(R)$ is in Laffey-Reams form, that is, $a_{ij} = 0$ for $j \geq i + 2$ and $A \equiv a_{11}1_n \pmod{(a_{12})}$. Then there exists an $\mathbf{x} = (x_1, \dots, x_{n-1})^\top \in R^{n-1}$, with $x_{n-1} = a_{11}$, such that*

$$\text{tr}(X(\mathbf{x}, a_{12})^r A) = 0,$$

for each $r = 1, \dots, n-1$.

Proof. By Lemma 3.2 we have

$$X(\mathbf{x}, a_{12})^r = \begin{pmatrix} 0 & \bar{0} \\ P^{r-1}\mathbf{x} & P^r \end{pmatrix},$$

where $P = (p_{ij})$, $1 \leq i, j \leq n-1$ is such that $p_{i,i+1} = 1$ for $i = 1, \dots, n-2$, $p_{n-1,1} = a_{12}$ and $p_{ij} = 0$ otherwise (i.e., P is the row-wise companion matrix of $x^{n-1} - a_{12}$). Writing A in block-form, we have

$$A = \begin{pmatrix} a_{11} & (a_{12}, 0, \dots, 0) \\ \mathbf{a} & Q \end{pmatrix},$$

where \mathbf{a} is an $n \times 1$ matrix and $Q \in \mathfrak{gl}_{n-1}(R)$. Thus

$$X(\mathbf{x}, a_{12})^r A = \begin{pmatrix} 0 & \bar{0} \\ a_{11}P^{r-1}\mathbf{x} + P^r\mathbf{a} & Q' \end{pmatrix},$$

where $Q' = P^{r-1}\mathbf{x}(a_{12}, 0, \dots, 0) + P^r Q$. Thus, by Lemma 3.1,

$$\text{tr}(X(\mathbf{x}, a_{12})^r A) = \text{tr}(Q') = a_{12}x_r + \text{tr}(P^r Q),$$

for each $r = 1, \dots, n-1$. We have $\text{tr}(P^r) \equiv 0 \pmod{(a_{12})}$, for $r = 1, \dots, n-1$, and since $A \equiv a_{11}1_n \pmod{(a_{12})}$ it follows that $Q \equiv a_{11}1_{n-1} \pmod{(a_{12})}$. Thus

$$\text{tr}(P^r Q) \equiv a_{11} \text{tr}(P^r) \equiv 0 \pmod{(a_{12})},$$

so there exist $m_r \in R$ such that $\text{tr}(P^r Q) = a_{12}m_r$, for each $r = 1, \dots, n-1$. Put $x_r = -m_r$, so that

$$\text{tr}(X(\mathbf{x}, a_{12})^r A) = 0,$$

for $r = 1, \dots, n-1$.

Finally, we claim that $\text{tr}(P^{n-1}Q) = -a_{11}a_{12}$, so that

$$x_{n-1} = a_{11}.$$

Indeed, since P is $\mathfrak{gl}_{n-1}(R)$ -regular with characteristic polynomial $x^{n-1} - a_{12}$, we have $P^{n-1} = a_{12}1_{n-1}$, so $\text{tr}(P^{n-1}Q) = a_{12}\text{tr}(Q) = a_{12}(-a_{11})$, as claimed. \square

The following result is essentially [9, Theorem 6.3], but the result here is stronger in that it says that X can be taken in $\mathfrak{sl}_n(R)$ and such that it is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular mod any maximal ideal \mathfrak{p} of R .

Theorem 5.2. *Let $A \in \mathfrak{sl}_n(R)$ with $n \geq 2$. Then there exist matrices $X \in \mathfrak{sl}_n(R)$ and $Y \in \mathfrak{gl}_n(R)$ such that $[X, Y] = A$, where X can be chosen such that $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular for every maximal ideal \mathfrak{p} of R .*

Proof. For $n = 2$ this is proved separately (see the proof of [9, Theorem 6.3]). Assume from now on that $n \geq 3$. First, if A is scalar, then $A \in \mathfrak{sl}_n(R)$ implies that either $A = 0$ or $n = 0$ in R . The former case is trivial, while the latter follows from Lemma 4.1.

Assume now that A is not scalar and let $A = (a_{ij})$. After a possible $\text{GL}_n(R)$ -conjugation, we can assume that A is in Laffey–Reams form; see [9, Theorem 5.6]. Moreover, we may assume that $(a_{11}, a_{12}) = (1)$, because if d is a common divisor of a_{11} and a_{12} , we can write $A = dA'$ for A' in Laffey–Reams form and if $A' = [X, Y]$ with X, Y as in the theorem, then $A = [X, dY]$.

By Lemma 5.1, there exists an $\mathbf{x} = (x_1, \dots, x_{n-1})^T \in R^{n-1}$, with $x_{n-1} = a_{11}$, such that

$$\text{tr}(X(\mathbf{x}, a_{12})^r A) = 0,$$

for each $r = 1, \dots, n-1$. Since $x_{n-1} = a_{11}$ and $(a_{11}, a_{12}) = (1)$, we have, for every maximal ideal \mathfrak{p} of R , that either $x_{n-1} \notin \mathfrak{p}$ or $a_{12} \notin \mathfrak{p}$, and therefore $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular, by Lemma 3.3. Thus, by [9, Proposition 3.3], there exists a $Y \in \mathfrak{gl}_n(R)$ such that

$$[X(\mathbf{x}, a_{12}), Y] = A.$$

\square

We now come to the proof of our main theorem. Just like the proof of the above theorem, our proof uses Lemma 5.1, but since here $X(\mathbf{x}, a_{12})_{\mathfrak{p}}$ cannot in general be $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular for all maximal ideals (cf. Remark 4.2), we need to treat the exceptional primes separately, and this requires us to pass to the localisations $R_{\mathfrak{p}}$, for various prime ideals $\mathfrak{p} \in \text{Spec}(R)$. For an element $X \in \mathfrak{gl}_n(R)$ we will write $X(\mathfrak{p})$ for its canonical image in $\mathfrak{gl}_n(R_{\mathfrak{p}})$, not to be confused with $X_{\mathfrak{p}} \in \mathfrak{gl}_n(R/\mathfrak{p})$. For any element $x \in R$, we will use the same symbol x to denote the image of x under the canonical injection $R \hookrightarrow R_{\mathfrak{p}}$, and the context will make it clear in which ring we are working. Similarly, we will denote the maximal ideal of $R_{\mathfrak{p}}$ by \mathfrak{p} and will identify $X_{\mathfrak{p}} \in \mathfrak{gl}_n(R/\mathfrak{p})$ with the image of $X(\mathfrak{p})$ in $\mathfrak{gl}_n(R_{\mathfrak{p}}/\mathfrak{p})$.

We will prove that for fixed $A, X \in \mathfrak{sl}_n(R)$, and for any maximal ideal \mathfrak{p} of R , there exists a solution $Y(\mathfrak{p}) \in \mathfrak{sl}_n(R_{\mathfrak{p}})$ to the localised equation $[X(\mathfrak{p}), Y(\mathfrak{p})] = A(\mathfrak{p})$. Since the equations $[X, Y] = A$, $\text{tr}(Y) = 0$ in Y are equivalent to a system of linear

equations in the entries of Y , the well known (and easy to prove) local-global principle for systems of linear equations (see, e.g., [3, Proposition 1]) implies the existence of a global solution.

Theorem 5.3. *Let $A \in \mathfrak{sl}_n(R)$ for $n \geq 3$. Then there exist matrices $X, Y \in \mathfrak{sl}_n(R)$ such that $[X, Y] = A$, where X can be chosen such that $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular for every maximal ideal \mathfrak{p} of R . Moreover, X can be chosen such that $X_{\mathfrak{p}}$ is $\mathfrak{sl}_n(R/\mathfrak{p})$ -regular for every \mathfrak{p} such that $A_{\mathfrak{p}}$ is not scalar.*

Proof. Assume first that A is scalar. Then $A \in \mathfrak{sl}_n(R)$ implies that either $A = 0$ or $n = 0$ in R . The former case is trivial, while the latter follows from Lemma 4.1.

Assume from now on that A is not scalar and let $A = (a_{ij})$. After a possible $\mathrm{GL}_n(R)$ -conjugation, we can assume that A is in Laffey–Reams form. Moreover, we may assume that $(a_{11}, a_{12}) = (1)$, because if d is a common divisor of a_{11} and a_{12} , we can write $A = dA'$ for A' in Laffey–Reams form, and if A' is a commutator of two matrices in $\mathfrak{sl}_n(R)$, then so is A .

By Lemma 5.1, there exists an $\mathbf{x} = (x_1, \dots, x_{n-1})^T \in R^{n-1}$, with $x_{n-1} = a_{11}$, such that

$$\mathrm{tr}(X(\mathbf{x}, a_{12})^r A) = 0,$$

for each $r = 1, \dots, n-1$. From now on, let $X := X(\mathbf{x}, a_{12})$. Since $(a_{11}, a_{12}) = (1)$, we have, for every maximal ideal \mathfrak{p} of R , that either $x_{n-1} \notin \mathfrak{p}$ or $a_{12} \notin \mathfrak{p}$, and therefore that $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular; see Lemma 3.3. Moreover, since A is in Laffey–Reams form, we have $A \equiv a_{11}1_n \pmod{(a_{12})}$, and this, combined with the fact that $\mathrm{tr}(A) = 0$ and $(a_{11}, a_{12}) = (1)$, implies that

$$(5.1) \quad n \in (a_{12}).$$

We will now pass to the localisations $R_{\mathfrak{p}}$ for various maximal ideals \mathfrak{p} of R . Let \mathfrak{p} be any maximal ideal of R . Then we have the local relation

$$\mathrm{tr}(X(\mathfrak{p})^r A(\mathfrak{p})) = 0, \quad r = 1, \dots, n-1.$$

in $R_{\mathfrak{p}}$. First, suppose that $A_{\mathfrak{p}}$ is not scalar. Then $a_{12} \notin \mathfrak{p}$, so the matrix $X(\mathfrak{p})_{\mathfrak{p}} = X_{\mathfrak{p}}$ is $\mathfrak{sl}_n(R_{\mathfrak{p}}/\mathfrak{p})$ -regular, by Lemma 3.3, and so, by Proposition 2.4, there exists a $Y(\mathfrak{p}) \in \mathfrak{sl}_n(R_{\mathfrak{p}})$ such that

$$[X(\mathfrak{p}), Y(\mathfrak{p})] = A(\mathfrak{p}).$$

Next, suppose that $A_{\mathfrak{p}}$ is scalar, so that $a_{12} \in \mathfrak{p}$. Since $a_{12} \neq 0$, X is $\mathfrak{sl}_n(F)$ -regular as an element of $\mathfrak{sl}_n(F)$, by Lemma 3.3, so there exists a $Y(0) \in \mathfrak{sl}_n(F)$ such that $[X, Y(0)] = A$. Clearing denominators in $Y(0)$ and passing to the localisation at \mathfrak{p} , we conclude that there exists a power p^m of a generator $p \in R_{\mathfrak{p}}$ of \mathfrak{p} and a $Q \in \mathfrak{sl}_n(R_{\mathfrak{p}})$, such that

$$(5.2) \quad [X(\mathfrak{p}), Q] = p^m A(\mathfrak{p}).$$

Let $m \geq 0$ be the minimal integer such that (5.2) holds for some $Q \in \mathfrak{sl}_n(R_{\mathfrak{p}})$. We will show that $m = 0$. For a contradiction, assume that $m \geq 1$. Reducing (5.2) mod \mathfrak{p} , we obtain $[X_{\mathfrak{p}}, Q_{\mathfrak{p}}] = 0$, so $Q_{\mathfrak{p}}$ commutes with $X_{\mathfrak{p}}$. Since $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(R/\mathfrak{p})$ -regular,

$$Q = f(X(\mathfrak{p})) + pD,$$

for some polynomial $f(T) \in R_{\mathfrak{p}}[T]$ of degree $n-1$ and some $D \in \mathfrak{gl}_n(R_{\mathfrak{p}})$. Write, $f(T) = c_0 + c_1T + \cdots + c_{n-1}T^{n-1}$, for $c_i \in R_{\mathfrak{p}}$. By Lemma (3.2), we have

$$\mathrm{tr}(X^i) = \begin{cases} n & \text{if } i = 0, \\ (n-1)a_{12} & \text{if } i = n-1, \\ 0 & \text{otherwise} \end{cases}$$

which implies

$$(5.3) \quad \mathrm{tr}(X(\mathfrak{p})^i) = \begin{cases} n & \text{if } i = 0, \\ (n-1)a_{12} & \text{if } i = n-1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$(5.4) \quad 0 = \mathrm{tr}(Q) = \sum_{i=0}^{n-1} c_i \mathrm{tr}(X(\mathfrak{p})^i) + p \mathrm{tr}(D) = c_0n + c_{n-1}(n-1)a_{12} + p \mathrm{tr}(D).$$

Moreover, we have $[X(\mathfrak{p}), Q] = [X(\mathfrak{p}), pD] = p^m A(\mathfrak{p})$, so

$$0 = \mathrm{tr}(pDp^m A(\mathfrak{p})) = p^{m+1} \mathrm{tr}(DA(\mathfrak{p})),$$

and thus $\mathrm{tr}(DA(\mathfrak{p})) = 0$. Since $A(\mathfrak{p}) \equiv a_{11}1_n \pmod{(a_{12})}$ and $(a_{11}, a_{12}) = (1)$, we conclude that

$$(5.5) \quad \mathrm{tr}(D) \in (a_{12}).$$

Since $n \in (a_{12})$ by (5.1), we have $n = a_{12}n'$ for some $n' \in R_{\mathfrak{p}}$. Moreover, since $R_{\mathfrak{p}}$ is a local ring, $n-1$ is a unit in $R_{\mathfrak{p}}$, so we can define the matrix

$$Q' = (c_0n'(n-1)^{-1} + c_{n-1})X^{n-1} + pD.$$

By (5.3) and (5.4) we have

$$\mathrm{tr}(Q') = c_0n + c_{n-1}(n-1)a_{12} + p \mathrm{tr}(D) = \mathrm{tr}(Q) = 0,$$

which, by (5.5), implies that $c_0n + c_{n-1}(n-1)a_{12} \in (pa_{12})$, and thus

$$c_0n'(n-1)^{-1} + c_{n-1} \in (p).$$

Writing $c_0n'(n-1)^{-1} + c_{n-1} = p\alpha$ for some $\alpha \in R_{\mathfrak{p}}$, we then get

$$[X(\mathfrak{p}), Q] = [X(\mathfrak{p}), pD] = [X(\mathfrak{p}), Q'] = p[X(\mathfrak{p}), \alpha X^{n-1} + D] = p^m A(\mathfrak{p}),$$

where $\mathrm{tr}(\alpha X^{n-1} + D) = 0$ because

$$p \mathrm{tr}(\alpha X^{n-1} + D) = \mathrm{tr}((c_0n'(n-1)^{-1} + c_{n-1})X^{n-1} + pD) = \mathrm{tr}(Q') = 0.$$

By cancelling a factor of p , we obtain a contradiction to the minimality of m in (5.2). Thus $m = 0$, so there exists a $Y(\mathfrak{p}) \in \mathfrak{sl}_n(R_{\mathfrak{p}})$ such that $[X(\mathfrak{p}), Y(\mathfrak{p})] = A(\mathfrak{p})$.

We have thus proved that for any maximal ideal \mathfrak{p} of R , there exists a $Y(\mathfrak{p}) \in \mathfrak{sl}_n(R_{\mathfrak{p}})$ such that

$$[X(\mathfrak{p}), Y(\mathfrak{p})] = A(\mathfrak{p}).$$

Thus, by the local-global principle for systems of linear equations (see, e.g., [3, Proposition 1]), there exists a $Y \in \mathfrak{sl}_n(R)$ such that

$$[X, Y] = A.$$

□

In the same way as in [9, Corollary 6.4], Theorem 5.3 implies the analogous statement over any principal ideal ring (PIR), thanks to a theorem of Hungerford that any PIR is a finite product of homomorphic images of PIDs.

6. SHALEV'S CONJECTURE FOR $n = 2$

This section is devoted to a proof of Shalev's conjecture mentioned in the introduction, in the case $n = 2$. For a group G and elements $x, y \in G$ we write the commutator as $(x, y) = xyx^{-1}y^{-1}$. In this section, R will denote a local PID (i.e., a discrete valuation ring) with residue field k . We denote the maximal ideal in R by \mathfrak{p} and let $\pi \in \mathfrak{p}$ be a generator. We first prove a result whose conclusion is weaker than Shalev's conjecture for $n = 2$, in that one of the elements is only shown to lie in $\mathrm{GL}_2(R)$, but where the hypotheses are slightly more general in that we allow any residue field apart from \mathbb{F}_2 . We will then refine this result to prove Proposition 6.2, which contains Shalev's conjecture for $n = 2$ as a special case.

Recall that an element $X \in \mathfrak{gl}_n(R)$ is $\mathfrak{gl}_n(R)$ -regular if and only if $X_{\mathfrak{p}}$ is $\mathfrak{gl}_n(k)$ -regular; see [9, Lemma 2.5]. Moreover, if X is $\mathfrak{gl}_n(R)$ -regular then X is $\mathrm{GL}_n(R)$ -conjugate to a companion matrix; see [9, Lemma 2.3]. By convention, we will write companion matrices row-wise, that is, with ones on the super-diagonal. If $X \in \mathfrak{gl}_n(R)$ has units on the superdiagonal and zeros above the superdiagonal, then it is regular; see [9, Lemma 2.7] (note that to go from units to ones on the superdiagonal, one only needs to conjugate by a diagonal element). The same conclusion holds if 'superdiagonal' is replaced by 'subdiagonal'.

Proposition 6.1. *Let R be a local PID such that $|k| > 2$. Then, for every $A \in \mathrm{SL}_2(R)$ there exist $x \in \mathrm{GL}_2(R)$ and $y \in \mathrm{SL}_2(R)$ such that $(x, y) = A$. Moreover, y can be taken to be $\mathrm{GL}_2(R)$ -conjugate to an element of the form $\begin{pmatrix} 0 & 1 \\ -1 & s \end{pmatrix} \in \mathrm{SL}_2(R)$.*

Proof. The strategy of the proof is the following. We find a $y \in \mathrm{SL}_2(R)$ such that both y and yA are regular as elements in $\mathfrak{gl}_n(R)$, and such that y and yA have the same determinant and trace. This implies that y and yA are $\mathrm{GL}_2(R)$ -conjugate, that is, there exists an $x \in \mathrm{GL}_2(R)$ such that

$$yA = xyx^{-1},$$

and so $(x, y) = A$.

By [9, Lemma 2.7], any

$$y = \begin{pmatrix} y_{11} & 1 \\ y_{21} & y_{22} \end{pmatrix} \in \mathfrak{gl}_n(R)$$

is $\mathfrak{gl}_n(R)$ -regular. In order for $\det(y) = \det(yA)$ we need $\det(y) = 1$, so $y_{21} = y_{11}y_{22} - 1$, which we assume henceforth. We distinguish two cases: either the image $A_{\mathfrak{p}}$ of A in $\mathrm{SL}_2(k)$ is scalar or regular.

Case 1: Assume that $A_{\mathfrak{p}}$ is scalar. If moreover A is scalar, then $A = \pm 1$, so $yA = \pm y$ is regular, and $\mathrm{tr}(yA) = \pm \mathrm{tr}(y)$. Setting $y_{22} = -y_{11}$ so that $\mathrm{tr}(y) = 0$, we obtain the existence of an $x \in \mathrm{GL}_2(R)$ such that $yA = xyx^{-1}$.

If $A_{\mathfrak{p}}$ is scalar but A is not scalar, then up to conjugation of A , we can write

$$A = \lambda 1 + \pi^i A',$$

where $\lambda = \pm 1$, $i \geq 1$ and $A' = \begin{pmatrix} 0 & 1 \\ a' & b' \end{pmatrix} \in \mathrm{SL}_2(R)$. Then, since $y_{\mathfrak{p}} \in \mathrm{SL}_2(k)$ is regular, and

$$(yA)_{\mathfrak{p}} = \pm y_{\mathfrak{p}},$$

we also have that $(yA)_{\mathfrak{p}}$ is regular, and hence (because R is local), that yA is regular. Furthermore, we need to ensure that

$$(6.1) \quad \mathrm{tr}(yA) = y_{11}(\pi^i y_{22} + \lambda) + y_{22}(b' \pi^i + \lambda) + \pi^i(a' - 1) = \mathrm{tr}(y) = y_{11} + y_{22}$$

If $\mathrm{char} k \neq 2$ and $\lambda = -1$, then $\lambda - 1$ is a unit, so this equation clearly has a solution $y_{11} \in R$ for any $y_{22} \in R$. If $\lambda = 1$ (6.1) is equivalent to

$$y_{11}y_{22} + y_{22}b' + a' - 1 = 0,$$

which has a solution over R , for example $y_{11} = -a' - b' + 1, y_{22} = 1$.

Case 2: Assume that $A_{\mathfrak{p}}$ is non-scalar. Then $A_{\mathfrak{p}}$ is regular, so A is regular, and hence, up to conjugation, we can write

$$A = \begin{pmatrix} 0 & 1 \\ -1 & b \end{pmatrix},$$

for $b \in R$. For $y = \begin{pmatrix} y_{11} & 1 \\ y_{11}y_{22} - 1 & y_{22} \end{pmatrix}$, we have

$$yA = \begin{pmatrix} -1 & y_{11} + b \\ -y_{22} & y_{22}(y_{11} + b) - 1 \end{pmatrix}.$$

Since the residue field of R has at least three elements, we can choose y_{22} such that both y_{22} and $y_{22} - 1$ are units. Then yA is regular by [9, Lemma 2.7], and the equation

$$\mathrm{tr}(yA) = y_{22}(y_{11} + b) - 2 = \mathrm{tr}(y) = y_{11} + y_{22},$$

has a solution $y_{11} \in R$. Hence, there exists an $x \in \mathrm{GL}_2(R)$ such that $ya = xyx^{-1}$.

Finally, $y = \begin{pmatrix} y_{11} & 1 \\ y_{11}y_{22} - 1 & y_{22} \end{pmatrix}$ is regular, so it is $\mathrm{GL}_2(R)$ -conjugate to $y = \begin{pmatrix} 0 & 1 \\ -1 & y_{11} + y_{22} \end{pmatrix}$. □

Note that the hypothesis on the residue field in the above proposition is optimal, since $\mathrm{GL}_2(\mathbb{F}_2) = \mathrm{SL}_2(\mathbb{F}_2)$, and this group is not perfect.

The following result implies Shalev's conjecture mentioned in the introduction, in the case $n = 2$.

Proposition 6.2. *Let R be a Henselian discrete valuation ring with finite residue field k such that $\mathrm{char} k > 2$ and $|k| > 3$. Then, for every $A \in \mathrm{SL}_2(R)$ there exist $x \in \mathrm{SL}_2(R)$ and $y \in \mathrm{SL}_2(R)$ such that $(x, y) = A$.*

Proof. If $A_{\mathfrak{p}}$ is non-scalar, the result follows from [2, Theorem 3.5]. Assume therefore that $A_{\mathfrak{p}}$ is scalar. By Proposition 6.1 we have $(x, y) = a$ for some $x \in \mathrm{GL}_2(R)$ and some y conjugate to $\begin{pmatrix} 0 & 1 \\ -1 & s \end{pmatrix}$, for $s \in R$. As before, let $y_{\mathfrak{p}} \in \mathrm{SL}_2(k)$ denote the image of y under the canonical map $\mathrm{SL}_2(R) \rightarrow \mathrm{SL}_2(k)$. We will show that s can be chosen such that $y_{\mathfrak{p}}$ is semisimple. If $y_{\mathfrak{p}}$ is not semisimple, then it has characteristic polynomial $x^2 - sx + 1 \equiv (x \pm 1)^2 \pmod{\mathfrak{p}}$, so $s \equiv \pm 2 \pmod{\mathfrak{p}}$. We therefore want to show that s can be chosen such that $s \not\equiv \pm 2 \pmod{\mathfrak{p}}$. We will examine Case 1 of the proof of Proposition 6.1 to show that this can indeed be achieved. As before, write

$$A = \lambda 1 + \pi^i A',$$

where $\lambda = \pm 1$ and $\begin{pmatrix} 0 & 1 \\ a' & b' \end{pmatrix}$. We have seen that if A is scalar, we can choose $s = 0 \not\equiv \pm 2 \pmod{\mathfrak{p}}$, and if $\lambda = -1$, (6.1) implies that $s = 0 \not\equiv \pm 2 \pmod{\mathfrak{p}}$. Moreover, if $\lambda = 1$, the last part of Case 1 of the proof of Proposition 6.1 says that we need to consider the relation

$$(6.2) \quad y_{11}y_{22} + y_{22}b' + a' - 1 = 0.$$

Note that $\det(A) = 1 + b'\pi - a'\pi^2 = 1$, so $b' \in \mathfrak{p}$. We consider the number of solutions of (6.2) mod \mathfrak{p} such that $s = y_{11} + y_{22} \equiv \pm 2 \pmod{\mathfrak{p}}$. The system of congruences

$$\begin{cases} (y_{11}y_{22} + a' - 1)_{\mathfrak{p}} &= 0, \\ (y_{11} + y_{22})_{\mathfrak{p}} &= \pm 2. \end{cases}$$

has at most 4 distinct solutions $((y_{11})_{\mathfrak{p}}, (y_{22})_{\mathfrak{p}}) \in k^2$ (at most two for each choice of sign for ± 2). On the other hand, the equation

$$(y_{11}y_{22} + a' - 1)_{\mathfrak{p}} = 0$$

has at least $|k| - 1$ distinct solutions, so when $|k| > 5$, there exist $(y_{11})_{\mathfrak{p}}, (y_{22})_{\mathfrak{p}} \in k$ such that $(y_{11} + y_{22})_{\mathfrak{p}} \not\equiv \pm 2$. Furthermore, if $k = \mathbb{F}_5$, the equation $(y_{11}y_{22} + a' - 1)_{\mathfrak{p}} = 0$ has 9 solutions if $a'_{\mathfrak{p}} = 1$, so it only remains to consider the case $k = \mathbb{F}_5$, with $a'_{\mathfrak{p}} \neq 1$. In this case, one checks easily that either $(y_{11})_{\mathfrak{p}} = 1, (y_{22})_{\mathfrak{p}} = 1 - a'$ or $(y_{11})_{\mathfrak{p}} = 2, (y_{22})_{\mathfrak{p}} = (1 - a')2^{-1}$ is a solution to $(y_{11}y_{22} + a' - 1)_{\mathfrak{p}} = 0$ such that $(y_{11} + y_{22})_{\mathfrak{p}} \not\equiv \pm 2$. We have therefore shown that whenever $|k| > 3$, there exist $y_{11}, y_{22} \in R$ such that $y_{11}y_{22} + a' - 1 \equiv 0 \pmod{\mathfrak{p}}$ and such that $s = y_{11} + y_{22} \not\equiv \pm 2 \pmod{\mathfrak{p}}$. Hensel's lemma now implies that equation (6.2) has a solution $y_{11}, y_{22} \in R$ such that $s = y_{11} + y_{22} \not\equiv \pm 2 \pmod{\mathfrak{p}}$. We thus conclude that $s = y_{11} + y_{22}$ can be chosen such that $y_{\mathfrak{p}}$ is semisimple.

Suppose, as we may, that $s \not\equiv \pm 2 \pmod{\mathfrak{p}}$, so that $y_{\mathfrak{p}}$ is semisimple. We claim that the determinant map

$$\det : C_{\mathrm{GL}_2(k)}(y_{\mathfrak{p}}) \longrightarrow k^{\times}$$

is surjective. Indeed, since $y_{\mathfrak{p}}$ is regular, we have

$$C_{\mathrm{GL}_2(k)}(y_{\mathfrak{p}}) = k[y_{\mathfrak{p}}]^{\times},$$

and if the characteristic polynomial of $y_{\mathfrak{p}}$ is irreducible, $k[y_{\mathfrak{p}}]$ is a field, while otherwise $y_{\mathfrak{p}}$ has distinct eigenvalues in k . When $k[y_{\mathfrak{p}}]$ is a field, the determinant coincides with the norm map $N : k[y_{\mathfrak{p}}]^{\times} \rightarrow k^{\times}$, which is well-known to be surjective (because k is finite). When $k[y_{\mathfrak{p}}]$ has distinct eigenvalues in k , the centraliser $C_{\mathrm{GL}_2(k)}(k[y_{\mathfrak{p}}])$ is $\mathrm{GL}_2(k)$ -conjugate to the diagonal subgroup of $\mathrm{GL}_2(k)$, on which \det is clearly surjective. We now want to show that

$$\det : C_{\mathrm{GL}_2(R)}(y) \longrightarrow R^{\times}$$

is surjective. Since y is regular, we have

$$C_{\mathfrak{gl}_2(R)}(y) = R[y] = \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha + \beta s \end{pmatrix} \mid \alpha, \beta \in R, \right\},$$

so the determinant is surjective if $\alpha^2 + \alpha\beta s + \beta^2 = r$ has a solution $\alpha, \beta \in R$ for each $r \in R^\times$. We show that Hensel's lemma implies that this equation has a solution $\alpha, \beta \in R$. Indeed, if the gradient

$$\begin{pmatrix} 2\alpha + \beta s \\ 2\beta + \alpha s \end{pmatrix} \equiv 0 \pmod{\mathfrak{p}},$$

then $\alpha \equiv \frac{\alpha s^2}{4}$, so either $\alpha \equiv 0$ or $s^2 \equiv 4$. In the same way, either $\beta \equiv 0$ or $s^2 \equiv 4$. Then, since we have assumed that $s \not\equiv \pm 2$, we obtain, $\alpha \equiv \beta \equiv 0$, which is not the case if $\alpha^2 + \alpha\beta s + \beta^2 = r \in R^\times$. Thus, any solution to this equation mod \mathfrak{p} has a lift to R . Since we have just observed that this equation has a solution mod \mathfrak{p} for every $r \in R^\times$, we conclude that $\det : C_{\mathrm{GL}_2(R)}(y) \rightarrow R^\times$ is surjective.

We have shown that there exists a $g \in C_{\mathrm{GL}_2(R)}(y)$ such that $\det(g) = \det(x)^{-1}$. Thus

$$yA = xyx^{-1} = (xg)y(xg)^{-1},$$

so $(xg, y) = A$ with $xg \in \mathrm{SL}_2(R)$. □

REFERENCES

- [1] A. A. Albert and B. Muckenhoupt, *On matrices of trace zero*, Michigan Math. J. **4** (1957), 1–3.
- [2] N. Avni, T. Gelander, M. Kassabov, and A. Shalev, *Word values in p -adic and adelic groups*, Bull. Lond. Math. Soc. **45** (2013), no. 6, 1323–1330.
- [3] J. A. Hermida and T. Sánchez-Giralda, *Linear equations over commutative rings and determinantal ideals*, J. Algebra **99** (1986), no. 1, 72–79.
- [4] T. J. Laffey and R. Reams, *Integral similarity and commutators of integral matrices*, Linear Algebra Appl. **197/198** (1994), 671–689.
- [5] D. Lissner, *Matrices over polynomial rings*, Trans. Amer. Math. Soc. **98** (1961), 285–305.
- [6] C. de Seguins Pazzis, *Commutators from a hyperplane of matrices*, Electron. J. Linear Algebra **27** (2014), 39–54.
- [7] A. Shalev, *Applications of some zeta functions in group theory*, in *Zeta functions in algebra and geometry*, volume 566 of *Contemp. Math.*, pp. 331–344, Amer. Math. Soc., Providence, RI, 2012.
- [8] A. Shalev, *Some results and problems in the theory of word maps*, in *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pp. 611–649, János Bolyai Math. Soc., Budapest, 2013.
- [9] A. Stasinski, *Similarity and commutators of matrices over principal ideal rings*, Trans. Amer. Math. Soc. **368** (2016), 2333–2354.
- [10] R. C. Thompson, *Matrices with zero trace*, Israel J. Math. **4** (1966), 33–42.
- [11] L. N. Vaserstein, *Noncommutative number theory*, in *Algebraic K-theory and Algebraic Number Theory*, volume 83 of *Contemp. Math.*, pp. 445–449, AMS, Providence, RI, 1989.

DEPARTMENT OF MATHEMATICAL SCIENCES, DURHAM UNIVERSITY, SOUTH RD, DURHAM, DH1 3LE, UK

E-mail address: alexander.stasinski@durham.ac.uk